

# Application and Prospect of Hospital Statistical Data Mining in the Network Environment

Xueyun Xia

Linshu County People's Hospital, Linyi, Shandong, 276700, China

## Abstract

As a new way to collect and process information, networks connect the logical and objective physical world together. In other words, it is a way for people to collect data and realize the computer world and the human social world, which has a huge impact on the collection of statistical data. If future statistical data collection can be integrated with the network to provide reliable data sources for industrial statistics, then the efficiency and data quality of statistical work will be improved. In order to further improve the level of hospital data mining, the author analyzes the hospital information system, expounds the mining and application of statistical information system, found that the deep use of data can further explore the law and the potential connection, data not only can provide strong support for the hospital leadership decision-making, but also can further improve the core competitiveness of the hospital.

## Keywords

network technology; data acquisition; application

## 网络环境下医院统计数据挖掘的应用与展望

夏雪筠

临沭县人民医院, 中国·山东临沂 276700

## 摘要

网络作为一种收集和處理信息的新方式, 将逻辑和客观的物理世界连接在一起。换句话说, 它是人们收集数据并实现计算机世界和人类社会世界的一种方式, 对统计数据的收集产生巨大影响。如果将来的统计数据收集可以与网络集成在一起, 为工业统计提供可靠的数据源, 那么统计工作的效率和数据质量将得到提高。而为了进一步提高医院的数据挖掘水平, 笔者通过对医院的信息系统进行分析, 阐述了统计信息系统的再次挖掘和应用, 发现数据的深层利用可以进一步挖掘出数据之间的规律以及存在的潜在联系, 不仅可以为医院领导层决策提供强有力的支撑, 同时还能进一步提高医院的核心竞争力。

## 关键词

网络技术; 数据采集; 应用

## 1 引言

网络是一种信息载体, 在医疗数据是非常实用的宝藏, 如患者的发病征兆状况与医生的处方及治疗效果等, 对医生的医术、医院的服务水平都有着相当大的促进作用。例如, 互联网或传统电信网络将所有简单、独立的功能进行互连互接。当前, 网络方法尚未广泛用于统计工作中, 统计工作仍然基于传统的工作方法, 没有实现自动化和信息化。传统的统计工作存在效率低, 存在成本高和误差大等许多问题。近期, 国家统计局积极促进了 PDA 数据采集技术的运用, 这极大地

提高了工作效率, 但是 PDA 数据采集技术还是基于手动操作, 这并没有开始改变传统的数据收集方式, 也没有利用当前快速发展的网络。例如, 云计算、人工智能等各种互联网先进技术。因此, 要想提高医院统计工作的质量, 科学合理地进行统计工作, 提高统计速度和可信性, 解决当前统计数据收集的难题, 重点加强对发现网络应用于统计数据采集的应用方法, 并尝试将当今先进的计算机网络和统计分析技术集合在一起, 实现数据自动化医院数据的目标, 通过建立网络技术于数据统计工作的关系, 能够大大提高当前的医院统计工作效率、准确性和便利性, 这在医院中非常重要, 如图 1 所示。

【作者简介】夏雪筠(1972-), 女, 中国山东临沂人, 本科, 高级会计师, 从事医院资源统计研究。

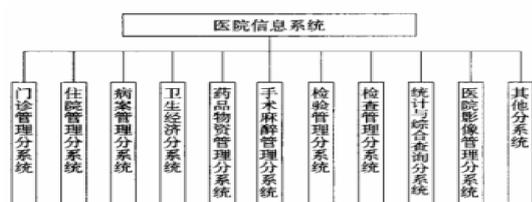


图1 医院信息系统的基本框架结构

## 2 医疗统计数据挖掘的特点

医疗数据挖掘十分独特，与其他挖掘领域有鲜明的区别。因为医疗数据是从情况不一的患者的诊断治疗中总结的，包括与患者的谈话纪录、诸多医学仪器拍摄的病情图片、医生主观的望闻问切等。医疗统计数据挖掘具有海量性，具体是医疗统计工作涉及的大量人员、临床数据等，导致医疗数据非常庞大；多种类别，指的是数值型数据、图片、文字、信号、视频和语音等，类别众多；模糊性，指的是专家们统计好医疗数据后分析融合了大量类别的数据时，相关人员不能快速准确地通过科技手段处理数据；主观性，具体是医生对患者的诊断有其主观性，患者对自身病情的把握也有主观性，这就意味着医疗数据整合的难度大；伦理性，具体是医学涉及的不仅是自然科学，也包括人文社科等，主要指数据归属权问题、数据安全问题、法律诉讼问题等<sup>[1]</sup>。

## 3 应用网络技术采集数据的技术路线

### 3.1 网络数据的采集

医院信息系统的应用，在一定程度上使医院的统计工作发生了巨大的转变，在开展传统医院统计工作的过程中，工作人员都是由单纯地完成报表任务以及理性分析，转变为现阶段较为先进的决策提供信息支撑的工作模式。通过使用先进的网络技术，例如射频技术、光电技术、二维码技术、无线网络技术、电子代码技术等。一方面，网络数据收集更方便，更快捷；另一方面，通过使用最新技术，可以接收到大量的数据，这可以有效改善传统的数据单一、数据丢失等缺点。

### 3.2 网络数据的整理

网络具有实时性、动态变化性、海量性、异构性等特性，为了应对缺失值和离群值，本研究采用了基于监控对象的建模，以最新的方法整理数据。

## 3.3 数据流的分类

数据清理后，网络数据显示为一系列真实、海量、动态变化的数据流，根据对数据流使用侧重点不同，此调查提供了一种对数据流进行分类的方法。根据概念漂移检测数据流分类的方法进行分类，使用不同的分类方法，用户可以针对每个侧重点获得不同的分类结果。

## 3.4 数据的统计分析

通过以下网络数据分析过程，本研究将现代统计技术与网络数据集成在一起，并克服了传统工作在数据收集，统计数据分析和整理方面的单一性和被动性缺点，让统计效率更高<sup>[1]</sup>。

## 4 网络技术在统计数据采集中的应用

### 4.1 扫描解码技术的应用

网络环境下大数据的挖掘在医疗方面意义重大。由于深入分析医疗统计的大数据，专家、学者甚至普通人都能在原始数据上提出自己的意见和建议，这些都是在大量数据的基础上完成的。条形码扫描是最常见的自动识别方法，识别系统通常由三部分组成。

首先，第一批扫描系统由光学组件和一个探测器组成，该阶段是二维码的扫描阶段，即将图案中的光信号换为相应的电信号。

其次，信号的整形步骤是转换电信号的步骤。经过放大，过滤形成后的矩形方波信号，然后对其进行译码。

最后，译码后的结果将发送到数据收集器。随着条码技术在商业中的广泛应用，本研究设计了更加良好的物价采价方案和相应的工具，以提高物价采价的效率。使用这些功能来获取物价采价方法特别是该系统的光电扫描系统被融入PDA中，以在商店中的后台销售系统中连接小装置，通过扫描产品的条形码，获取价格信息<sup>[2]</sup>。

### 4.2 RFID 扫描解码技术的应用

RFID技术具有识别和读取相对长距离信息的能力，因此该电子产品还具有很高的存储容量，从而更轻松有效地帮助企业进行商品管理，还可以使得销售企业与制造公司的关系密切，两家公司之间的协作以及促进沟通和交流，加快对信息的反馈，从而优化了生产和销售链，另外可以使用RFID

技术获取价格信息,相关带着RFID读取器的PDA,并检索指定位置的规格品,扫一下就可以将产品信息录入,录入价格后,信息将自动发送到统计部门。PDA分三步使用。

#### 4.2.1 信息收集层

此层通过采价员和PDA来完成。如上所述,此PDA由RFID系统,条形码扫描系统和CPI数据收集软件组成,采价员可以通过这多功能的PDA来获取商品价格信息<sup>[3]</sup>。

#### 4.2.2 信息处理层

信息处理层是一个平台,用于处理收集的数据并支持信息的输入,处理和传输。该集成平台增加了第三方中间件应用程序界面,使其与PDA采集系统软件兼容。该软件主要用于数据处理,如数据清理和数据挖掘。

#### 4.2.3 信息应用层

此级别通过信息使用者完成,当数据经过处理层分类后,然后将信息传输到管理者那,管理者通过对数据进行分析以了解当前价格变化。

### 4.3 文本扫描技术的应用

文本扫描(OCR)技术通常称为光学字符识别技术。此方法首先通过扫描(或光学输入)将文本信息录入,然后将文本转换为图像(更准确地说是图像信息),然后将图像信息将转换为易于使用文字,可以发送此文本信息,最后由PDA或计算机将识别结果显示出来。换句话说,此方法使可以扫描特定图像或材料的数据并对其进行处理,并生成新文本信息,最后被人们所使用。文本扫描技术进行价格采集的工作中,最重要的技术是OCR技术,OCR技术已经成熟,可以广泛使用。另外,网络正在引起广泛的共识和关注,政府大力支持和推广,因此将互联网用于政府物价统计是会得到支持的。系统硬件要求是开发OCR软件客户端和扫描仪,从而最大程度地降低了相关数据库的成本。OCR技术的引入可以显著减小劳动强度,节省大量的人工成本<sup>[4]</sup>。

### 4.4 音频解码技术的应用

音频解码方法也称为语音识别技术。目前,语音识别技

术基本上实现了自由说语音识别,在算法模型方面实现了快速发展,语音技术正在工业中得到应用,尤其是电子产品这一领域,语音识别技术被广泛用于该领域实现其功能。语音识别技术中最重要的事情是语音模式的安装和语音信号特征对比,在市场环境中,由于具有各种噪声,所以生成语音模式更加困难。这里采用隐马尔可夫模型,并且使用各种方法来实现语音识别<sup>[5]</sup>。

## 5 结语

基于网络技术收集统计数据的主要好处之一是实时数据收集系统,该系统可显著提高数据收集效率并达到较高水平,提高数据质量。为了促进数据统计技术的使用,相关部门机构必须建立统一标准,根据不同级别的需求和统计部门的统计数据收集,指引企业建立数据采集平台,以便于国家统计局各种业务系统的集成构建和使用。医学数据仓库是数据仓库技术面向医学领域数据的具体实现,与其他企业数据仓库相比,医学数据仓库的数据来源、数据类型和数据特征都有其特殊之处。建立医学数据仓库是医学数据分析处理的基础,是医学信息技术发展的必然,对于医学、医疗卫生、药理学和医学管理等领域的研究与应用都有巨大的推动作用。

## 参考文献

- [1] 任栋,董雪建,曹改改,等. 网络技术在统计数据采集中的应用探索[J]. 调研世界,2020(4):62-65.
- [2] 于秋红. 网络技术在大型公共建筑能耗数据采集中的应用[J]. 现代电子技术,2018,41(11):139-143.
- [3] 辛金国,龚恺. 基于网络技术统计数据采集的影响因素[J]. 统计与决策,2015(23):34-37.
- [4] 辛金国,龚恺,王渊. 基于网络技术的统计数据采集系统构建探索[J]. 统计与咨询,2013(6):31-32.
- [5] 何登峰. 基于数据挖掘技术的计算机辅助概念设计的研究[D]. 镇江:江苏科技大学,2004.